



# On Infinite Real Trace Rational Languages of Maximum Topological Complexity

Olivier Finkel, Jean-Pierre Ressayre, Pierre Simonnet

## ► To cite this version:

Olivier Finkel, Jean-Pierre Ressayre, Pierre Simonnet. On Infinite Real Trace Rational Languages of Maximum Topological Complexity. *Zapiski Nauchnyh Seminarov POMI*, 2004, 316, pp.205-223. hal-00109918v2

**HAL Id: hal-00109918**

**<https://hal.science/hal-00109918v2>**

Submitted on 3 Jan 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# On infinite real trace rational languages of maximum topological complexity \*

<sup>1</sup>Olivier Finkel and <sup>2</sup>Jean-Pierre Ressayre and <sup>3</sup>Pierre Simonnet

<sup>1,2</sup>*Equipe de Logique Mathématique*

U.F.R. de Mathématiques, Université Paris 7

2 Place Jussieu 75251 Paris cedex 05, France.

finkel@logique.jussieu.fr   ressayre@logique.jussieu.fr

<sup>3</sup>*UMR CNRS 6134, Faculté des Sciences, Université de Corse*

Quartier Grossetti BP52 20250, Corte, France

simonnet@univ-corse.fr

## Abstract

We consider the set  $\mathbb{R}^\omega(\Gamma, D)$  of infinite real traces, over a dependence alphabet  $(\Gamma, D)$  with no isolated letter, equipped with the topology induced by the prefix metric. We then prove that all rational languages of infinite real traces are analytic sets. We reprove also that there exist some rational languages of infinite real traces which are analytic but non Borel sets, and even  $\Sigma_1^1$ -complete, hence of maximum possible topological complexity. For that purpose we give an example of  $\Sigma_1^1$ -complete language which is fundamentally different from the known example of  $\Sigma_1^1$ -complete infinitary rational relation given in [Fin03c].

**Keywords:** Real traces; rational languages; topological properties; analytic and Borel sets.

## 1 Introduction

Trace monoids were firstly considered by Cartier and Foata for studying combinatorial problems, [CF69]. Next Mazurkiewicz introduced finite traces as a semantic model for concurrent systems, [Maz77]. Since then traces have been much investigated by various authors and they have been extended to infinite traces to model systems which may not terminate; see the handbook [DR95] and its chapter about infinite traces [GP95], for many results and references.

In particular, **real traces** have been studied by Gastin, Petit and Zielonka, who characterized in [GPZ94] the two important families of recognizable and rational languages of real traces, over a dependence alphabet  $(\Gamma, D)$ , in connection with rational languages of finite or infinite words.

Several metrics have been defined on the set  $\mathbb{R}(\Gamma, D)$  of real traces over  $(\Gamma, D)$ . In

---

\*The results of this paper have been exposed during the International Conference JAF 22, 22nd Journées sur les Arithmétiques Faibles, June 11-14, 2003, Napoli, Italy.

particular, the prefix metric defined by Kwiatkowska [Kwi90] and the Foata normal form metric defined by Bonnizzoni, Mauri and Pighizzini [BMP90]. Kummeretz and Kuske stated in [KK03] that for *finite* dependence alphabets these two metrics define the same topology on  $\mathbb{R}(\Gamma, D)$ . Moreover, if we consider only **infinite real traces** over a dependence alphabet  $(\Gamma, D)$  without isolated letter, the topological subspace  $\mathbb{R}^\omega(\Gamma, D) = \mathbb{R}(\Gamma, D) - \mathbb{M}(\Gamma, D)$  of  $\mathbb{R}(\Gamma, D)$  (where  $\mathbb{M}(\Gamma, D)$  is the set of finite traces over  $(\Gamma, D)$ ), is homeomorphic to the Cantor set  $2^\omega$ , or equivalently to any set  $\Sigma^\omega$  of infinite words over a finite alphabet  $\Sigma$ , equipped with the product of the discrete topology on  $\Sigma$ , [KK03, Sta97a, PP04]. We can then define, from open subsets of the topological space  $\mathbb{R}^\omega(\Gamma, D)$ , the hierarchy of Borel sets by successive operations of countable intersections and countable unions. Furthermore, it is well known that there exist some subsets of the Cantor set, hence also some subsets of  $\mathbb{R}^\omega(\Gamma, D)$ , which are not Borel. There is another hierarchy beyond the Borel one, called the projective hierarchy.

It is then natural to try to locate classical languages of infinite real traces with regard to these hierarchies and this question is posed by Lescow and Thomas in [LT94] (for the general case of infinite labelled partial orders like traces). In the case of infinite words, Mc Naughton's Theorem implies that every  $\omega$ -regular language is a boolean combination of  $\Pi_2^0$ -sets hence a  $\Delta_3^0 = (\Pi_3^0 \cap \Pi_3^0)$ -set. Landweber studied first the topological properties of  $\omega$ -regular languages and characterized the  $\omega$ -regular languages in each of the Borel classes  $\Sigma_1^0$ ,  $\Pi_1^0$ ,  $\Sigma_2^0$ ,  $\Pi_2^0$ , [Lan69]. We study in this paper the topological complexity of **rational languages of infinite real traces**.

We show below that all rational languages of infinite real traces are analytic sets and that there exist some rational languages of infinite real traces which are analytic but non Borel sets, and even  $\Sigma_1^1$ -complete, hence of maximum possible topological complexity, giving a partial answer to the question of the comparison between the topological complexity of rational languages of infinite words and of infinite traces [LT94].

The first author recently showed in [Fin03c] that there exists a  $\Sigma_1^1$ -complete infinitary rational relation  $R \subseteq \Sigma_1^\omega \times \Sigma_2^\omega$  where  $\Sigma_1$  and  $\Sigma_2$  are two finite alphabets having at least two letters.

We could have used this result to prove that there exists a  $\Sigma_1^1$ -complete rational language of infinite real traces  $L \subseteq \mathbb{R}^\omega(\Gamma, D)$ , whenever  $(\Gamma, D)$  is a dependence alphabet and  $\Gamma \supseteq \Sigma_1 \cup \Sigma_2$ , where  $\Sigma_1$  and  $\Sigma_2$  are two independent dependence cliques having at least two letters. This can be done by considering the natural embedding  $i : \Sigma_1^\omega \times \Sigma_2^\omega \rightarrow \mathbb{R}^\omega(\Gamma, D)$ . The language  $R' = i(R)$  is then  $\Sigma_1^1$ -complete. But this way the language  $R'$  would have in fact the structure of an infinitary rational relation.

On the other side the  $\Sigma_1^1$ -complete language  $\mathcal{L}$  given in this paper is a new example whose structure is radically different from that of  $R'$ . In particular,  $\mathcal{L}$  does not contain any  $\Sigma_1^1$ -complete language of infinite traces having the structure of an infinitary rational relation.

This is important because in Trace Theory the structure of dependence alphabets is very important: some results are known to be true for some dependence alphabets and false for other dependence alphabets (see for example [HM97]).

Moreover we think that the presentation of this new example has also some

interest for the following reasons. The proof given in this paper is self contained and is stated in the general context of traces. The problem is exposed in this general context and we use here the general properties of traces instead of the particular properties of infinitary rational relations. We use the characterization of rational languages of infinite real traces given by Theorem 2.1 of Gastin, Petit and Zielonka, which states a connection between rational languages of infinite words and rational languages of infinite real traces instead of the notion of Büchi transducer.

We prove also that all rational languages of infinite real traces are analytic. Our proof is not difficult, but it is original, using the Baire space  $\omega^\omega$  and the characterization of rational languages of infinite real traces given by Theorem 2.1: every rational language of infinite traces is a finite union of sets of the form  $R.S^\omega$  where  $S$  and  $R$  are rational monoalphabetic languages of finite traces. This proof is rather different from usual ones in the theory of  $\omega$ -languages. It uses a connection between the topological complexity of  $\omega$ -powers of languages of finite traces, i.e. of languages of the form  $S^\omega$ , where  $S$  is a language of finite traces, and the topological complexity of rational languages of infinite traces. The closure under countable union of the class of analytic sets is also important in the proof.

We expect that the inverse way could also be fruitful: it seems to us that there should exist some context free  $\omega$ -languages and some infinitary rational relations of high transfinite Borel rank. We think that we could use this fact to show that there exist  $\omega$ -powers of finitary languages of high transfinite Borel rank. Notice that the question of the topological complexity of  $\omega$ -powers (of languages of finite words) has been raised by several authors [Niw90, Sim92, Sta97a, Sta97b] and some new results have been recently proved [Fin01, Fin03a, Fin04b, Lec01].

The paper is organized as follows. In section 2 we recall the notion of words and traces. In section 3 we recall definitions of Borel and analytic sets, and we prove our main results in section 4.

## 2 Words and traces

Let us now introduce notations for words. For  $\Sigma$  a finite alphabet, a non empty finite word over  $\Sigma$  is a finite sequence of letters:  $x = a_1 a_2 \dots a_n$  where  $\forall i \in [1; n]$   $a_i \in \Sigma$ . We shall denote  $x(i) = a_i$  the  $i^{th}$  letter of  $x$  and  $x[i] = x(1) \dots x(i)$  for  $i \leq n$ . The length of  $x$  is  $|x| = n$ . The empty word will be denoted by  $\varepsilon$  and has no letters. Its length is 0. The set of non empty finite words over  $\Sigma$  is denoted  $\Sigma^+$ .  $\Sigma^* = \Sigma^+ \cup \{\varepsilon\}$  is the set of finite words over  $\Sigma$ . A (finitary) language  $L$  over  $\Sigma$  is a subset of  $\Sigma^*$ . The usual concatenation product of  $u$  and  $v$  will be denoted by  $u.v$  or just  $uv$ . For  $V \subseteq \Sigma^*$ , we denote  $V^* = \{v_1 \dots v_n \mid n \geq 1 \text{ and } \forall i \in [1; n] \ v_i \in V\} \cup \{\varepsilon\}$ .

The first infinite ordinal is  $\omega$ . An  $\omega$ -word over  $\Sigma$  is an  $\omega$ -sequence  $a_1 a_2 \dots a_n \dots$ , where  $\forall i \geq 1 \ a_i \in \Sigma$ . When  $\sigma$  is an  $\omega$ -word over  $\Sigma$ , we write  $\sigma = \sigma(1)\sigma(2) \dots \sigma(n) \dots$  and  $\sigma[n] = \sigma(1)\sigma(2) \dots \sigma(n)$  the finite word of length  $n$ , prefix of  $\sigma$ . The set of  $\omega$ -words over the alphabet  $\Sigma$  is denoted by  $\Sigma^\omega$ . An  $\omega$ -language over an alphabet  $\Sigma$  is a subset of  $\Sigma^\omega$ . For  $V \subseteq \Sigma^*$ ,  $V^\omega = \{\sigma = u_1 \dots u_n \dots \in \Sigma^\omega \mid \forall i \geq 1 \ u_i \in V\}$

is the  $\omega$ -power of  $V$ . The concatenation product is extended to the product of a finite word  $u$  and an  $\omega$ -word  $v$ : the infinite word  $u.v$  is then the  $\omega$ -word such that:  $(u.v)(k) = u(k)$  if  $k \leq |u|$ , and  $(u.v)(k) = v(k - |u|)$  if  $k > |u|$ .

The prefix relation is denoted  $\sqsubseteq$ : the finite word  $u$  is a prefix of the finite word  $v$  (respectively, the infinite word  $v$ ), denoted  $u \sqsubseteq v$ , if and only if there exists a finite word  $w$  (respectively, an infinite word  $w$ ), such that  $v = u.w$ .

We shall denote  $\Sigma^\infty = \Sigma^* \cup \Sigma^\omega$  the set of finite or infinite words over  $\Sigma$ .

We introduce firstly traces as dependence graphs, [DR95, GP95, KK03]. A dependence relation over an alphabet  $\Gamma$  is a reflexive and symmetric relation on  $\Gamma$ . Its complement  $I_D = (\Gamma \times \Gamma) - D$  is the independence relation induced by the relation  $D$ ; the relation  $I_D$  is irreflexive and symmetric. A dependence alphabet  $(\Gamma, D)$  is formed by a finite alphabet  $\Gamma$  and a dependence relation  $D \subseteq \Gamma \times \Gamma$ . A dependence graph  $[V, E, \lambda]$  over the dependence alphabet  $(\Gamma, D)$  is an isomorphism class of a node labelled graph  $(V, E, \lambda)$  such that  $(V, E)$  is a directed acyclic graph,  $V$  is at most countably infinite,  $\lambda : V \rightarrow \Gamma$  is a function which associates a label  $\lambda(a)$  to each node  $a \in V$ , and such that:

- (1)  $\forall v, w \in V \quad (\lambda(v), \lambda(w)) \in D \leftrightarrow (v = w \text{ or } (v, w) \in E \text{ or } (w, v) \in E)$
- (2) The reflexive and transitive closure  $E^*$  of the edge relation  $E$  is well founded, i.e. there is no infinite strictly decreasing sequence of vertices.

Let us remark that, since in this definition  $(V, E)$  is acyclic,  $E^*$  is a partial order on  $V$ .

The empty trace has no vertice and will be denoted by  $\varepsilon$  as in the case of words.

As usually the concatenation of two dependence graphs  $g_1 = [V_1, E_1, \lambda_1]$  and  $g_2 = [V_2, E_2, \lambda_2]$ , where we can assume, without loss of generality, that  $V_1$  and  $V_2$  are disjoint, is the dependence graph  $g_1.g_2 = [V, E, \lambda]$  such that  $V = V_1 \cup V_2$ ,  $E = E_1 \cup E_2 \cup \{(v_1, v_2) \in V_1 \times V_2 \mid (\lambda_1(v_1), \lambda_2(v_2)) \in D\}$ , and  $\lambda = \lambda_1 \cup \lambda_2$ .

The alphabet  $alph(t)$  of a trace  $t = [V, E, \lambda]$  is the set  $\lambda(V)$ . The alphabet at infinity of  $t$  is the set  $alphinf(t) = \{a \in \Gamma \mid \lambda^{-1}[a] \text{ is infinite}\}$  of all  $a \in alph(t)$  occurring infinitely often in  $t$ .

The set  $\mathbb{M}(\Gamma, D)$  of finite traces over  $(\Gamma, D)$  is the set of traces having only finitely many vertices. For  $t \in \mathbb{M}(\Gamma, D)$  the length of  $t$  is the number of vertices of  $t$  denoted  $|t|$ .

The star operation  $T \rightarrow T^*$  and the operation  $T \rightarrow T^\omega$  are naturally extended to subsets  $T$  of  $\mathbb{M}(\Gamma, D)$ :

$$T^* = \{t_1.t_2 \dots t_n \mid n \geq 1 \text{ and } \forall i \in [1; n] \ t_i \in T\} \cup \{\varepsilon\}$$

$$T^\omega = \{t_1.t_2 \dots t_n \dots \mid \forall i \ t_i \in T\}.$$

A real trace over  $(\Gamma, D)$  is a dependence graph  $[V, E, \lambda]$  such that for all  $v \in V$  the set  $\{u \in V \mid (u, v) \in E^*\}$  is finite. The set of real traces over  $(\Gamma, D)$  is denoted  $\mathbb{R}(\Gamma, D)$  and the set  $\mathbb{R}(\Gamma, D) - \mathbb{M}(\Gamma, D)$  of **infinite** real traces will be denoted by  $\mathbb{R}^\omega(\Gamma, D)$ .

The prefix order over words can be extended to real traces in the following way. For all  $s, t \in \mathbb{R}(\Gamma, D)$   $s \sqsubseteq t$  iff there exists  $z \in \mathbb{R}(\Gamma, D)$  such that  $s.z = t$  iff  $s$  is a downwards closed subgraph of  $t$ . The corresponding suffix  $z$  is then unique.

Real traces may also be viewed as equivalence classes of (finite or infinite) words. Let  $(\Gamma, D)$  be a dependence alphabet and let  $\varphi : \Gamma^\infty \rightarrow \mathbb{R}(\Gamma, D)$  be the mapping defined by  $\varphi(a) = [\{x\}, \emptyset, x \rightarrow a]$  for each  $a \in \Gamma$  and  $\varphi(a_1.a_2 \dots) = \varphi(a_1).\varphi(a_2) \dots$  for each  $a_1.a_2 \dots$  in  $\Gamma^\infty$ . Let us remark that if there is some  $(a, b)$  in  $I_D$  then the mapping is not injective because for instance  $\varphi(ab) = \varphi(ba)$ . One can define an equivalence relation  $\sim_I$  on  $\Gamma^\infty$  by: for all  $u, v \in \Gamma^\infty$   $u \sim_I v$  iff  $\varphi(u) = \varphi(v)$ . Then  $\varphi$  induces a surjective morphism from the free monoid  $\Gamma^*$  onto the monoid of finite dependence graphs  $\mathbb{M}(\Gamma, D) = \Gamma^* / \sim$ . And the set  $\varphi(\Gamma^\infty) = \Gamma^\infty / \sim$  is the set of real traces  $\mathbb{R}(\Gamma, D)$ .

The empty trace is the image  $\varphi(\varepsilon)$  of the empty word and is still denoted by  $\varepsilon$ .

We assume the reader to be familiar with the theory of formal languages and of  $\omega$ -regular languages, see [Tho90, Sta97a, PP04] for many results and references. We recall that  $\omega$ -regular languages are accepted by Büchi automata and that the class of  $\omega$ -regular languages is the omega Kleene closure of the class of regular finitary languages.

The family of rational real trace languages over  $(\Gamma, D)$  is the smallest family which contains the emptyset, all the singletons  $\{[\{x\}, \emptyset, x \rightarrow a]\}$ , for  $a \in \Gamma$ , and which is closed under finite union, concatenation product,  $\star$ -iteration and  $\omega$ -iteration on real traces.

Let us now recall the following characterization of rational languages of infinite real traces (there exists also a version for finite *or* infinite traces, [GPZ94]). A real trace language  $R$  is said to be monoalphabetic if  $\text{alph}(s) = \text{alph}(t)$  for all  $s, t \in R$ .

**Theorem 2.1 ([GPZ94])** *Let  $T \subseteq \mathbb{R}^\omega(\Gamma, D)$  be a language of infinite real traces over the dependence alphabet  $(\Gamma, D)$ . The following assertions are equivalent:*

- (1)  *$T$  is rational.*
- (2)  *$T$  is a finite union of sets of the form  $R.S^\omega$  where  $S$  and  $R$  are rational monoalphabetic languages of finite traces over  $(\Gamma, D)$  and  $\varepsilon \notin S$ .*
- (3)  *$T = \varphi(L)$  for some  $\omega$ -regular language  $L \subseteq \Gamma^\omega$ .*

### 3 Topology

We assume the reader to be familiar with basic notions of topology which may be found in [Kec95, LT94, Sta97a, PP04].

There is a natural metric on the set  $\Sigma^\omega$  of infinite words over a finite alphabet  $\Sigma$  which is called the prefix metric and defined as follows. For  $u, v \in \Sigma^\omega$  and  $u \neq v$  let  $d(u, v) = 2^{-l_{\text{pref}}(u, v)}$  where  $l_{\text{pref}}(u, v)$  is the first integer  $n$  such that the  $(n+1)^{\text{th}}$  letter of  $u$  is different from the  $(n+1)^{\text{th}}$  letter of  $v$ . This metric induces on  $\Sigma^\omega$  the usual Cantor topology for which open subsets of  $\Sigma^\omega$  are in the form  $W.\Sigma^\omega$ , where  $W \subseteq \Sigma^*$ .

(Notice that this prefix metric may be extended to the set  $\Sigma^\infty$  of finite or infinite words over the alphabet  $\Sigma$ ).

The prefix metric has been extended to real traces by Kwiatkowska in [Kwi90] by defining firstly for all  $s, t \in \mathbb{R}(\Gamma, D)$  with  $s \neq t$ :

$$l_{pref}(s, t) = \sup\{n \in \mathbb{N} \mid r \sqsubseteq s \leftrightarrow r \sqsubseteq t \text{ for all } r \in \mathbb{M}(\Gamma, D) \text{ with } |r| \leq n\}$$

and next  $d_{pref}(s, t) = 2^{-l_{pref}(s, t)}$

Notice that we consider in this paper *infinite* traces and Kwiatkowska defined the prefix metric over *finite or infinite* traces as one could also have done in the case of words. If  $D = \Gamma \times \Gamma$  the prefix metric on infinite real traces over  $(\Gamma, D)$  coincide with the preceding definition in the case of infinite words over  $\Gamma$ .

If  $(\Gamma, D)$  is a dependence alphabet, a letter  $a \in \Gamma$  is said to be an isolated letter if  $a$  is independent from all other letters of  $\Gamma$ , i.e.  $\forall b \in \Gamma - \{a\}, (a, b) \in I_D$ . From now on we suppose that a dependence alphabet has no isolated letter. Then the set  $\mathbb{R}^\omega(\Gamma, D)$  of infinite real traces over  $(\Gamma, D)$ , equipped with the topology induced by the prefix metric, is homeomorphic to the Cantor set  $\{0, 1\}^\omega$ , hence also to  $\Sigma^\omega$  for every finite alphabet  $\Sigma$  having at least two letters [KK03].

Borel subsets of the Cantor set (hence also of topological spaces  $\Sigma^\omega$  or  $\mathbb{R}^\omega(\Gamma, D)$ ) form a strict infinite hierarchy, the Borel hierarchy, which is defined from open sets by successive operations of countable unions and of countable intersections. We give the definition in the case of a topological space  $\Sigma^\omega$ , the definition being similar in the case of the topological space  $\mathbb{R}^\omega(\Gamma, D)$ . Then we recall some well known properties of Borel sets.

**Definition 3.1** *The classes  $\Sigma_n^0$  and  $\Pi_n^0$  of the Borel Hierarchy on the topological space  $\Sigma^\omega$  are defined as follows:*

$\Sigma_1^0$  *is the class of open subsets of  $\Sigma^\omega$ .*

$\Pi_1^0$  *is the class of closed subsets, i.e. complements of open subsets, of  $\Sigma^\omega$ .*

*And for any integer  $n \geq 1$ :*

$\Sigma_{n+1}^0$  *is the class of countable unions of  $\Pi_n^0$ -subsets of  $\Sigma^\omega$ .*

$\Pi_{n+1}^0$  *is the class of countable intersections of  $\Sigma_n^0$ -subsets of  $\Sigma^\omega$ .*

*The Borel Hierarchy is also defined for transfinite levels. The classes  $\Sigma_\alpha^0$  and  $\Pi_\alpha^0$ , for a non-null countable ordinal  $\alpha$ , are defined in the following way:*

$\Sigma_\alpha^0$  *is the class of countable unions of subsets of  $\Sigma^\omega$  in  $\cup_{\gamma < \alpha} \Pi_\gamma^0$ .*

$\Pi_\alpha^0$  *is the class of countable intersections of subsets of  $\Sigma^\omega$  in  $\cup_{\gamma < \alpha} \Sigma_\gamma^0$ .*

### Theorem 3.2

- (a)  $\Sigma_\alpha^0 \cup \Pi_\alpha^0 \subsetneq \Sigma_{\alpha+1}^0 \cap \Pi_{\alpha+1}^0$ , for each countable ordinal  $\alpha \geq 1$ .
- (b)  $\cup_{\gamma < \alpha} \Sigma_\gamma^0 = \cup_{\gamma < \alpha} \Pi_\gamma^0 \subsetneq \Sigma_\alpha^0 \cap \Pi_\alpha^0$ , for each countable limit ordinal  $\alpha$ .
- (c) A set  $W \subseteq \Sigma^\omega$  is in the class  $\Sigma_\alpha^0$  iff its complement is in the class  $\Pi_\alpha^0$ .
- (d)  $\Sigma_\alpha^0 - \Pi_\alpha^0 \neq \emptyset$  and  $\Pi_\alpha^0 - \Sigma_\alpha^0 \neq \emptyset$  hold for every countable ordinal  $\alpha \geq 1$ .

We shall say that a subset of  $\Sigma^\omega$  is a Borel set of rank  $\alpha$ , for a countable ordinal  $\alpha$ , iff it is in  $\Sigma_\alpha^0 \cup \Pi_\alpha^0$  but not in  $\bigcup_{\gamma < \alpha} (\Sigma_\gamma^0 \cup \Pi_\gamma^0)$ .

Let us recall the characterization of  $\Pi_2^0$ -subsets of  $\Sigma^\omega$ , involving the  $\delta$ -limit  $W^\delta$  of a finitary language  $W$ . For  $W \subseteq \Sigma^*$  and  $\sigma \in \Sigma^\omega$ ,  $\sigma \in W^\delta$  iff  $\sigma$  has infinitely many prefixes in  $W$ , i.e.  $W^\delta = \{\sigma \in \Sigma^\omega / \exists^\omega i \text{ such that } \sigma[i] \in W\}$ , see [Sta97a].

**Proposition 3.3** *A subset  $L$  of  $\Sigma^\omega$  is a  $\Pi_2^0$ -subset of  $\Sigma^\omega$  iff there exists a set  $W \subseteq \Sigma^*$  such that  $L = W^\delta$ .*

**Example 3.4** *Let  $\Sigma = \{0, 1\}$  and  $\mathcal{A} = (0^*.1)^\omega \subseteq \Sigma^\omega$  be the set of  $\omega$ -words over the alphabet  $\Sigma$  with infinitely many occurrences of the letter 1. It is well known that  $\mathcal{A}$  is a  $\Pi_2^0$ -subset of  $\Sigma^\omega$  because  $\mathcal{A} = ((0^*.1)^+)^{\delta}$  holds.*

There are some subsets of the Cantor set, (hence also of the topological spaces  $\Sigma^\omega$  or  $\mathbb{R}^\omega(\Gamma, D)$ ) which are not Borel sets. There exists another hierarchy beyond the Borel hierarchy, called the projective hierarchy. Projective sets are defined from Borel sets by successive operations of projection and complementation. We shall only need in this paper the first class of the projective hierarchy: the class  $\Sigma_1^1$  of **analytic** sets. A set  $A \subseteq \Sigma^\omega$  is analytic iff there exists a Borel set  $B \subseteq (\Sigma \times Y)^\omega$ , with  $Y$  a finite alphabet, such that  $x \in A \leftrightarrow \exists y \in Y^\omega$  such that  $(x, y) \in B$ , where  $(x, y) \in (\Sigma \times Y)^\omega$  is defined by:  $(x, y)(i) = (x(i), y(i))$  for all integers  $i \geq 1$ .

Analytic sets are also characterized as continuous images of the Baire space  $\omega^\omega$ , which is the set of infinite sequences of non negative integers. It may be seen as the set of infinite words over the infinite alphabet  $\omega = \{0, 1, 2, \dots\}$ . The topology of the Baire space is then defined by a prefix metric which is just an extension of the previous one to the case of an infinite alphabet.

A set  $A \subseteq \Sigma^\omega$  (respectively  $A \subseteq \mathbb{R}^\omega(\Gamma, D)$ ) is then analytic iff there exists a continuous function  $f : \omega^\omega \rightarrow \Sigma^\omega$  (respectively  $f : \omega^\omega \rightarrow \mathbb{R}^\omega(\Gamma, D)$ ) such that  $f(\omega^\omega) = A$ .

A  $\Sigma_\alpha^0$  (respectively  $\Pi_\alpha^0, \Sigma_1^1$ )-complete set is a  $\Sigma_\alpha^0$  (respectively  $\Pi_\alpha^0, \Sigma_1^1$ )- set which is in some sense a set of the highest topological complexity among the  $\Sigma_\alpha^0$  (respectively  $\Pi_\alpha^0, \Sigma_1^1$ )- sets. This notion is defined via reductions by continuous functions. More precisely a set  $F \subseteq \Sigma^\omega$  is said to be a  $\Sigma_\alpha^0$  (respectively  $\Pi_\alpha^0, \Sigma_1^1$ )-complete set iff for any set  $E \subseteq Y^\omega$  (with  $Y$  a finite alphabet):  $E \in \Sigma_\alpha^0$  (respectively  $E \in \Pi_\alpha^0, \Sigma_1^1$ ) iff there exists a continuous function  $f$  such that  $E = f^{-1}(F)$ .  $\Sigma_n^0$  (respectively  $\Pi_n^0$ )-complete sets, with  $n$  an integer  $\geq 1$ , are thoroughly characterized in [Sta86].

The  $\omega$ -regular language  $\mathcal{A} = (0^*.1)^\omega$  given in Example 3.4 is a well known example of  $\Pi_2^0$ -complete set.

## 4 Rational languages of infinite traces

We want now to investigate the topological complexity of rational languages of infinite real traces. In a first step we shall give an upper bound of this complexity, showing that all rational languages  $T \subseteq \mathbb{R}^\omega(\Gamma, D)$  are analytic sets.



We would like to use the characterization of rational languages  $T \subseteq \mathbb{R}^\omega(\Gamma, D)$  given in item 3 of Theorem 2.1:  $T = \varphi(L)$  for some  $\omega$ -regular language  $L \subseteq \Gamma^\omega$ . Indeed every  $\omega$ -regular language is a Borel set (of rank at most 3) and the continuous image of a Borel set is an analytic set. Unfortunately, the mapping  $\varphi$  is not continuous as the following example shows. Let  $(a, b) \in I_D$  and  $x_n \in \Gamma^\omega$  defined by  $x_n = a^n b a^\omega$  for each integer  $n \geq 1$ . Then in  $\Gamma^\omega$  the sequence  $(x_n)_{n \geq 1}$  is convergent and its limit is  $a^\omega$ . But the sequence  $(\varphi(x_n))_{n \geq 1}$  is constant in  $\mathbb{R}^\omega(\Gamma, D)$  because for all  $n \geq 1$   $\varphi(x_n) = \varphi(b a^\omega)$ . Thus the sequence  $(\varphi(x_n))_{n \geq 1}$  is convergent but its limit is  $\varphi(b a^\omega)$  which is different from  $\varphi(a^\omega)$ .

We shall use the characterization of rational languages  $T \subseteq \mathbb{R}^\omega(\Gamma, D)$  given in item 2 of Theorem 2.1:  $T$  is a finite union of sets of the form  $R.S^\omega$  where  $S$  and  $R$  are rational monoalphabetic languages of finite traces over  $(\Gamma, D)$  and  $\varepsilon \notin S$ . We consider firstly such rational languages in the simple form  $S^\omega$  where  $S$  is a monoalphabetic language of finite traces over  $(\Gamma, D)$  which does not contain the empty trace. The set  $S$  is at most countable so it can be finite or countably infinite. In the first case  $\text{card}(S)=p$  and we can fix an enumeration of  $S$  by a bijective function  $\psi : \{0, 1, 2, \dots, p-1\} \rightarrow S$  and in the second case we can fix an enumeration of  $S$  by a bijective function  $\psi : \omega = \{0, 1, 2, \dots\} \rightarrow S$ . Let now  $H$  be the function defined from  $\{0, 1, 2, \dots, p-1\}^\omega$  (in the first case) or from  $\omega^\omega$  (in the second case) into  $\mathbb{R}^\omega(\Gamma, D)$  by:

$$H(n_1 n_2 \dots n_i \dots) = \psi(n_1). \psi(n_2) \dots \psi(n_i) \dots$$

for all sequences  $n_1 n_2 \dots n_i \dots$  in  $\{0, 1, 2, \dots, p-1\}^\omega$  (in the first case) or in  $\omega^\omega$  (in the second case). It holds that  $H(\{0, 1, 2, \dots, p-1\}^\omega) = S^\omega$  (in the first case) or that  $H(\omega^\omega) = S^\omega$  (in the second case).

It is easy to see that  $H$  is a continuous function. A crucial point is that  $S$  is a **monoalphabetic** language, i.e. there exists  $\Gamma' \subseteq \Gamma$  such that for all  $s \in S$ ,  $\text{alph}(s) = \Gamma'$ . Let  $\mathcal{N} = (n_i)_{i \geq 1}$  and  $\mathcal{M} = (m_i)_{i \geq 1}$  be two infinite sequences of integers in  $\{0, 1, 2, \dots, p-1\}^\omega$  or in  $\omega^\omega$  such that for all  $i \leq k$   $n_i = m_i$ . Then  $r \sqsubseteq H(\mathcal{N}) \leftrightarrow r \sqsubseteq H(\mathcal{M})$  holds (at least) for all  $r \in \mathbb{M}(\Gamma, D)$  with  $|r| \leq k$ . Thus  $l_{\text{pref}(H(\mathcal{N}), H(\mathcal{M}))} \geq k$  and  $d_{\text{pref}}(H(\mathcal{N}), H(\mathcal{M})) = 2^{-l_{\text{pref}(H(\mathcal{N}), H(\mathcal{M}))}} \leq 2^{-k}$ . This implies that the function  $H$  is continuous (and even uniformly continuous).

If  $S$  is finite, then the set  $S^\omega$  is the continuous image of the compact set  $\{0, 1, 2, \dots, p-1\}^\omega$  thus it is a closed hence also an analytic subset of  $\mathbb{R}^\omega(\Gamma, D)$ . If  $S$  is infinite, then the set  $S^\omega$  is the continuous image of the Baire space  $\omega^\omega$  thus it is an analytic set.

Let now  $R \subseteq \mathbb{M}(\Gamma, D)$  be a language of finite traces. For  $r \in R$  let  $\theta_r : \mathbb{R}^\omega(\Gamma, D) \rightarrow \mathbb{R}^\omega(\Gamma, D)$  be the function defined by  $\theta_r(t) = r.t$  for all  $t \in \mathbb{R}^\omega(\Gamma, D)$ . It is easy to see that this function is continuous. Then  $r.S^\omega = \theta_r(S^\omega)$  is an analytic set because the image of an analytic set by a continuous function is still an analytic set. The language  $R.S^\omega = \bigcup_{r \in R} r.S^\omega$  is a countable union of analytic sets (because  $R$  is countable) but the class of analytic subsets of  $\mathbb{R}^\omega(\Gamma, D)$  is closed under countable unions thus  $R.S^\omega$  is an analytic set.

A rational language  $T \subseteq \mathbb{R}^\omega(\Gamma, D)$  is a finite union of sets of the form  $R.S^\omega$  where  $S$  and  $R$  are rational monoalphabetic languages of finite traces over  $(\Gamma, D)$  and  $\varepsilon \notin S$ . Then by finite union this language is an analytic set.

Notice that we have not used the fact that  $S$  and  $R$  are rational so the above proof can be applied to finite unions of sets of the form  $R.S^\omega$  where  $S$  is a monoalphabetic language of finite traces and we have got the following result.

**Proposition 4.1** *Let  $(\Gamma, D)$  be a dependence alphabet without isolated letter, and let  $S_i, R_i, 1 \leq i \leq n$ , be languages of finite traces over  $(\Gamma, D)$ , where, for all  $i$ ,  $S_i$  does not contain the empty trace and is monoalphabetic. Then the language of infinite traces*

$$T = \bigcup_{1 \leq i \leq n} R_i.S_i^\omega$$

*is an analytic set. In particular every rational language  $T \subseteq \mathbb{R}^\omega(\Gamma, D)$  is an analytic set.*

In order to prove the existence of  $\Sigma_1^1$ -complete rational language of infinite traces, we shall use results about languages of infinite binary trees whose nodes are labelled in a finite alphabet  $\Sigma$  having at least two letters.

A node of an infinite binary tree is represented by a finite word over the alphabet  $\{l, r\}$  where  $r$  means “right” and  $l$  means “left”. Then an infinite binary tree whose nodes are labelled in  $\Sigma$  may be viewed as a function  $t : \{l, r\}^* \rightarrow \Sigma$ . The set of infinite binary trees labelled in  $\Sigma$  will be denoted  $T_\Sigma^\omega$ .

There is a natural topology on this set  $T_\Sigma^\omega$  which is defined by the following distance, [LT94]. Let  $t$  and  $s$  be two distinct infinite trees in  $T_\Sigma^\omega$ . Then the distance between  $t$  and  $s$  is  $\frac{1}{2^n}$  where  $n$  is the smallest integer such that  $t(x) \neq s(x)$  for some word  $x \in \{l, r\}^*$  of length  $n$ .

The open sets are then in the form  $T_0.T_\Sigma^\omega$  where  $T_0$  is a set of finite labelled trees.  $T_0.T_\Sigma^\omega$  is the set of infinite binary trees which extend some finite labelled binary tree  $t_0 \in T_0$ ,  $t_0$  is here a sort of prefix, an “initial subtree” of a tree in  $T_\Sigma^\omega$ .

It is well known that the topological space  $T_\Sigma^\omega$  is homeomorphic to the Cantor set hence also to the topological spaces  $\Sigma^\omega$  or  $\mathbb{R}^\omega(\Gamma, D)$ .

The Borel hierarchy and the projective hierarchy on  $T_\Sigma^\omega$  are defined from open sets as in the cases of the topological spaces  $\Sigma^\omega$  or  $\mathbb{R}^\omega(\Gamma, D)$ .

Let  $t$  be a tree. A branch  $B$  of  $t$  is a subset of the set of nodes of  $t$  which is linearly ordered by the tree partial order  $\sqsubseteq$  and which is closed under prefix relation, i.e. if  $x$  and  $y$  are nodes of  $t$  such that  $y \in B$  and  $x \sqsubseteq y$  then  $x \in B$ . A branch  $B$  of a tree is said to be maximal iff there is no other branch of  $t$  which strictly contains  $B$ .

Let  $t$  be an infinite binary tree in  $T_\Sigma^\omega$ . If  $B$  is a maximal branch of  $t$ , then this branch is infinite. Let  $(u_i)_{i \geq 0}$  be the enumeration of the nodes in  $B$  which is strictly increasing for the prefix order.

The infinite sequence of labels of the nodes of such a maximal branch  $B$ , i.e.  $t(u_0)t(u_1)\dots t(u_n)\dots$  is called a path. It is an  $\omega$ -word over the alphabet  $\Sigma$ .

For  $L \subseteq \Sigma^\omega$  we denote  $Path(L)$  the set of infinite trees  $t$  in  $T_\Sigma^\omega$  such that  $t$  has at least one path in  $L$ .

It is well known that if  $L \subseteq \Sigma^\omega$  is a  $\Pi_2^0$ -complete subset of  $\Sigma^\omega$  (or a Borel set of higher complexity in the Borel hierarchy) then the set  $Path(L)$  is a  $\Sigma_1^1$ -complete subset of  $T_\Sigma^\omega$ , [Niw85, Sim93], [PP04, exercise].

In order to use this result we shall firstly code trees labelled in  $\Sigma$  by infinite words over the finite alphabet  $\Gamma = \Sigma \cup \Sigma' \cup \{A, B\}$  where  $\Sigma' = \{a' \mid a \in \Sigma\}$  is a disjoint copy of the alphabet  $\Sigma$  and  $A, B$  are additional letters not in  $\Sigma \cup \Sigma'$ .

Consider now the set  $\{l, r\}^*$  of nodes of binary infinite trees. For each integer  $n \geq 0$ , call  $C_n$  the set of words of length  $n$  of  $\{l, r\}^*$ . Then  $C_0 = \{\varepsilon\}$ ,  $C_1 = \{l, r\}$ ,  $C_2 = \{ll, lr, rl, rr\}$  and so on.  $C_n$  is the set of nodes which appear at the  $(n+1)^{th}$  level of an infinite binary tree. The number of nodes of  $C_n$  is  $card(C_n) = 2^n$ . We consider now the lexicographic order on  $C_n$  (assuming that  $l$  is before  $r$  for this order). Then, in the enumeration of the nodes with regard to this order, the nodes of  $C_1$  will be:  $l, r$ ; the nodes of  $C_3$  will be:  $lll, llr, lrl, lrr, rll, rlr, rrl, rrr$ . Let  $u_1^n, \dots, u_j^n, \dots, u_{2^n}^n$  be such an enumeration of  $C_n$  in the lexicographic order and let  $v_1^n, \dots, v_j^n, \dots, v_{2^n}^n$  be the enumeration of the elements of  $C_n$  in the reverse order. Then for all integers  $n \geq 0$  and  $i$ ,  $1 \leq i \leq 2^n$ , it holds that  $v_i^n = u_{2^n+1-i}^n$ .

For  $t \in T_\Sigma^\omega$  let  $U_n^t = t(u_1^n)t(u_2^n) \dots t(u_{2^n}^n)$  be the finite word enumerating the labels of nodes in  $C_n$  in the lexicographic order, and let  $V_n^t = t(v_1^n)t(v_2^n) \dots t(v_{2^n}^n)$  be the reverse sequence. Let  $V_n'^t = \psi(V_n^t)$  where  $\psi$  is the morphism from  $\Sigma^*$  into  $\Sigma'^*$  defined by  $\psi(a) = a'$  for all  $a \in \Sigma$ . The code  $g(t)$  of  $t$  is then

$$g(t) = V_0'^t . A . U_1^t . B . V_2'^t . A . U_3^t . B . V_4'^t . A \dots A . U_{2n+1}^t . B . V_{2n+2}'^t . A \dots$$

The  $\omega$ -word  $g(t)$  enumerates the labels of the nodes of the tree  $t$  which appear at successive levels  $1, 2, 3, \dots$ . The (images by  $\psi$  of) labels of nodes occurring at odd level  $2n+1$  are enumerated in the reverse lexicographic order by the sequence  $V_{2n}'^t$  and the labels of nodes occurring at even level  $2n$  are enumerated in the lexicographic order by the sequence  $U_{2n-1}^t$ . Labels of nodes of distinct levels are alternatively separated by a letter  $A$  or a letter  $B$ .

Let now  $(\Gamma, D)$  be a dependence alphabet where  $\Gamma = \Sigma \cup \Sigma' \cup \{A, B\}$  and the independence relation  $I_D = \Gamma \times \Gamma - D$  is defined by

$$I_D = \Sigma \times (\{A\} \cup \Sigma') \bigcup (\{A\} \cup \Sigma') \times \Sigma \bigcup \Sigma' \times \{B\} \bigcup \{B\} \times \Sigma'$$

i.e. letters of  $\Sigma$  may only commute with  $A$  and with letters in  $\Sigma'$  while letters of  $\Sigma'$  may only commute with  $B$  and with letters in  $\Sigma$ , letter  $A$  (respectively  $B$ ) may only commute with letters in  $\Sigma$  (respectively, with letters in  $\Sigma'$ ).

Let now  $h : T_\Sigma^\omega \rightarrow \mathbb{R}^\omega(\Gamma, D)$  be the function defined by:

$$\forall t \in T_\Sigma^\omega \quad h(t) = \varphi(g(t))$$

We firstly state the following result.

**Lemma 4.2** *The above defined function  $h : T_\Sigma^\omega \rightarrow \mathbb{R}^\omega(\Gamma, D)$  is continuous.*

**Proof.** Let us remark that in a segment

$$B.V_{2n}'^t.A.U_{2n+1}^t.B.V_{2n+2}'^t.A$$

of an  $\omega$ -word  $g(t)$  written as above, letters of  $U_{2n+1}^t$  may only commute with the preceding letter  $A$  and letters of  $V_{2n}'^t$ . In a similar manner letters of  $V_{2n+2}'^t$  may only commute with the preceding letter  $B$  and letters of  $U_{2n+1}^t$ .

So if two infinite binary trees  $t, s \in T_\Sigma^\omega$  have the same labels on their  $k$  first levels, ( $k > 1$ ), then for all  $r \in \mathbb{M}(\Gamma, D)$  such that

$$|r| \leq (k-1) + 1 + 2 + 2^2 + \dots + 2^{k-2}$$

it holds that

$$r \sqsubseteq h(t) \leftrightarrow r \sqsubseteq h(s)$$

So

$$l_{pref(h(t), h(s))} \geq (k-1) + 1 + 2 + 2^2 + \dots + 2^{k-2} \geq 2^{k-1}$$

and

$$d_{pref}((h(t), h(s))) = 2^{-l_{pref}(h(t), h(s))} \leq 2^{-2^{k-1}}$$

so we have proved:

$$\forall t, s \in T_\Sigma^\omega \quad d(t, s) \leq 2^{-k} \rightarrow d_{pref}((h(t), h(s))) \leq 2^{-2^{k-1}}$$

Thus the function  $h$  is continuous (and even uniformly continuous).  $\square$

Let now  $\mathcal{R} \subseteq \Sigma^\omega$  be a regular  $\omega$ -language. We are going to define from  $\mathcal{R}$  a language of infinite real traces  $\mathcal{L}$  over the dependence alphabet  $(\Gamma, D)$  defined above. Then we shall prove that  $\mathcal{L}$  is rational and that  $Path(\mathcal{R}) = h^{-1}(\mathcal{L})$ .

Let us firstly define  $\mathcal{L}$  as being the set of infinite traces  $\varphi(\sigma)$  where  $\sigma \in \Gamma^\omega$  may be written in the following form:

$$\sigma = x(1).u_1.A.v_1.x(2).u_2.B.v_2.x(3).u_3.A \dots A.v_{2n+1}.x(2n+2).u_{2n+2}.B.v_{2n+2}.x(2n+3).u_{2n+3}.A \dots$$

where for all integers  $i \geq 0$ ,

$$\begin{aligned} x(2i+2) &\in \Sigma & \text{and} & & x(2i+1) &\in \Sigma' \\ u_{2i+2}, v_{2i+1} &\in \Sigma^* & \text{and} & & u_{2i+1}, v_{2i+2} &\in \Sigma'^* \\ |v_i| &= 2|u_i| & \text{or} & & |v_i| &= 2|u_i| + 1 \end{aligned}$$

and the  $\omega$ -word

$$x = \psi^{-1}(x(1))x(2)\psi^{-1}(x(3)) \dots x(2n)\psi^{-1}(x(2n+1))x(2n+2) \dots \text{ is in } \mathcal{R}.$$

**Lemma 4.3** *The above defined language  $\mathcal{L}$  of infinite real traces is rational.*

**Proof.** Every  $\omega$ -word

$$\sigma = x(1).u_1.A.v_1.x(2).u_2.B.v_2.x(3).u_3.A \dots A.v_{2n+1}.x(2n+2).u_{2n+2}.B.v_{2n+2}.x(2n+3).u_{2n+3}.A \dots$$

written as above is equivalent, modulo the equivalence relation  $\sim_{I_D}$  over infinite words in  $\Gamma^\omega$ , to the infinite word

$$\sigma' = x(1).u_1.v_1.A.x(2).u_2.v_2.B.x(3).u_3.v_3.A \dots A.x(2n+2).u_{2n+2}.v_{2n+2}.B.x(2n+3).u_{2n+3}.v_{2n+3}.A \dots$$

But letters of  $\Sigma$  may commute also with letters of  $\Sigma'$  and for all integers  $i$ ,

$$|v_i| = 2|u_i| \quad \text{or} \quad |v_i| = 2|u_i| + 1$$

by definition of  $\mathcal{L}$ . Thus every  $\omega$ -word  $\sigma$  written as above is also equivalent, modulo  $\sim_{I_D}$ , to an infinite word in  $\Gamma^\omega$  in the form

$$\sigma'' = x(1).W_1.A.x(2).W_2.B.x(3).W_3.A \dots A.x(2n+2).W_{2n+2}.B.x(2n+3).W_{2n+3}.A \dots$$

where for all integers  $i \geq 0$ ,

$$W_{2i+1} \in (\Sigma' \Sigma^2)^* \cdot (\Sigma \cup \{\varepsilon\})$$

$$W_{2i+2} \in (\Sigma \Sigma'^2)^* \cdot (\Sigma' \cup \{\varepsilon\})$$

Let  $L$  be the  $\omega$ -language over the alphabet  $\Gamma$  formed by all such  $\omega$ -words  $\sigma''$  such that  $x = \psi^{-1}(x(1))x(2)\psi^{-1}(x(3)) \dots x(2n)\psi^{-1}(x(2n+1))x(2n+2) \dots$  is in  $\mathcal{R}$ .

It is easy to see that  $L$  is an  $\omega$ -regular language. Moreover  $\mathcal{L} = \varphi(L)$  thus we can infer from Theorem 2.1 that  $\mathcal{L}$  is a rational language of infinite real traces.  $\square$

We are going now to prove the following result.

**Lemma 4.4** *For  $\mathcal{L}$  defined as above from the  $\omega$ -language  $\mathcal{R}$ , it holds that  $\text{Path}(\mathcal{R}) = h^{-1}(\mathcal{L})$ , i.e.  $\forall t \in T_\Sigma^\omega \quad h(t) \in \mathcal{L} \iff t \in \text{Path}(\mathcal{R})$ .*

**Proof.** Suppose that  $h(t) \in \mathcal{L}$  for some  $t \in T_\Sigma^\omega$ . Then  $h(t) = \varphi(g(t)) = \varphi(\sigma)$  where  $\sigma \in \Gamma^\omega$  may be written in the following form:

$$\sigma = x(1).u_1.A.v_1.x(2).u_2.B.v_2.x(3).u_3.A \dots A.v_{2n+1}.x(2n+2).u_{2n+2}.B.v_{2n+2}.x(2n+3).u_{2n+3}.A \dots$$

where for all integers  $i \geq 0$ ,

$$x(2i+2) \in \Sigma \quad \text{and} \quad x(2i+1) \in \Sigma'$$

$$u_{2i+2}, v_{2i+1} \in \Sigma^* \quad \text{and} \quad u_{2i+1}, v_{2i+2} \in \Sigma'^*$$

$$|v_i| = 2|u_i| \quad \text{or} \quad |v_i| = 2|u_i| + 1$$

and the  $\omega$ -word

$$x = \psi^{-1}(x(1))x(2)\psi^{-1}(x(3)) \dots x(2n)\psi^{-1}(x(2n+1))x(2n+2) \dots \text{ is in } \mathcal{R}.$$

Then it is easy to see that  $\sigma = g(t)$ , because of the definition of  $\sigma$ , of  $g(t)$ , and of the independence relation  $I_D$  on  $\Gamma$ .

Then  $\psi^{-1}(x(1)) = t(v_1^0)$  and  $u_1 = \varepsilon$ , then  $|v_1| = 2|u_1| = 0$  or  $|v_1| = 2|u_1| + 1 = 1$ . If  $|v_1| = 0$  then  $x(2) = t(u_1^1)$  and if  $|v_1| = 1$  then  $x(2) = t(u_2^1)$ . Then the choice of  $|v_1| = 2|u_1|$  or of  $|v_1| = 2|u_1| + 1$  implies that  $x(2)$  is the label of the left or the right successor of the root node  $v_1^0 = \varepsilon$ .

This phenomenon will happen for next levels. The choice of  $|v_i| = 2|u_i|$  or of  $|v_i| = 2|u_i| + 1$  determines one of the two successor nodes of a node at level  $i$  (whose label is  $x(i)$  if  $i$  is even or  $\psi^{-1}(x(i))$  if  $i$  is odd) and then the label of this successor is  $\psi^{-1}(x(i+1))$  if  $i$  is even, or  $x(i+1)$  if  $i$  is odd.

Thus the successive choices determine a branch of  $t$  and the labels of nodes of this branch (changing only  $x(2n+1)$  in  $\psi^{-1}(x(2n+1))$ ) form a path  $x = \psi^{-1}(x(1))x(2)\psi^{-1}(x(3)) \dots x(2n)\psi^{-1}(x(2n+1))x(2n+2) \dots$  which is in  $\mathcal{R}$ . Then  $t \in \text{Path}(\mathcal{R})$ .

Conversely it is easy to see that if  $t \in \text{Path}(\mathcal{R})$ , the infinite word  $g(t)$  may be written as a word  $\sigma$  in the above form. Then  $h(t) = \varphi(g(t)) = \varphi(\sigma)$  is in  $\mathcal{L}$ .  $\square$

We can now state the following

**Theorem 4.5** *There exist some  $\Sigma_1^1$ -complete, hence non Borel, rational languages of infinite real traces.*

**Proof.** Suppose  $\mathcal{R} \subseteq \Sigma^\omega$  is a  $\Pi_2^0$ -complete  $\omega$ -regular language. Let then  $\mathcal{L} \subseteq \mathbb{R}^\omega(\Gamma, D)$  be defined as above.  $\mathcal{L}$  is a rational language of infinite real traces by Lemma 4.3. Then  $\mathcal{L}$  is an analytic subset of  $\mathbb{R}^\omega(\Gamma, D)$  by Proposition 4.1.

But  $\text{Path}(\mathcal{R})$  is a  $\Sigma_1^1$ -complete set and  $\text{Path}(\mathcal{R}) = h^{-1}(\mathcal{L})$  holds by Lemma 4.4 thus  $\mathcal{L}$  is also  $\Sigma_1^1$ -complete. In particular  $\mathcal{L}$  is not a Borel set.  $\square$

## 5 Concluding remarks

The existence of a  $\Sigma_1^1$ -complete infinitary rational relation has been used to get many undecidability results in [Fin03d] and the existence of a  $\Sigma_1^1$ -complete context free  $\omega$ -language led to other undecidability results in [Fin03a, Fin03b, FS03]. In particular, the topological complexity and the degree of ambiguity of an infinitary rational relation or of a context free  $\omega$ -language are highly undecidable.

In a similar way, the existence of a  $\Sigma_1^1$ -complete recognizable language of infinite pictures, proved in [ATW02] by Altenbernd, Thomas, and Wöhrle, has been used in [Fin04a] to prove many undecidability results, giving in particular the answer to some open questions of [ATW02].

Topological arguments following from the existence of  $\Sigma_1^1$ -complete rational languages of infinite real traces can also be used to prove similar undecidability results for languages of infinite traces.

In [FS03] have been established some links between the existence of a  $\Sigma_1^1$ -complete  $\omega$ -language in the form  $V^\omega$  and the number of decompositions of  $\omega$ -words of  $V^\omega$  in words of  $V$ .

We think that such facts could be useful in the domain of combinatorics of

traces. The code problem for traces is important in Trace Theory and several questions are still opened [HM97]. The analogue of the notion of  $\omega$ -code and the study of the number of decompositions of infinite traces of  $V^\omega$ , where  $V$  is a set of finite traces, in infinite product of traces of  $V$ , is also an important subject related to practical applications and to the notion of ambiguity (see [Aug01, AA01] for related results in the case of words). We think that topological arguments could be useful in this research area, and the existence of several  $\Sigma_1^1$ -complete languages of infinite traces, having different structures, could be useful in the cases of different dependence alphabets.

**Acknowledgements.** Thanks to the anonymous referees for useful comments on a preliminary version of this paper.

## References

- [Aug01] X. Augros, Des Algorithmes Autour des Codes Rationnels, Ph.D. Thesis, universit  de Nice-Soplia Antipolis, December 2001.
- [AA01] M. K. Ahmad and X. Augros, Some Results on Codes for Generalized Factorizations, Journal of Automata, Languages and Combinatorics, Volume 6 (3), p. 239-251, 2001.
- [ATW02] J-H. Altenbernd, W. Thomas and S. W hrle, Tiling Systems over Infinite Pictures and their Acceptance Conditions, in the Proceedings of the 6th International Conference on Developements in Language Theory, DLT 2002, Lecture Notes in Computer Science, Volume 2450, p. 297-306, Springer, 2003.
- [BMP90] P. Bonnizzoni, G. Mauri and G. Pighizzini, About Infinite Traces, in V. Diekert, editor, Proceedings of the ASMICS Workshop Free Partially Commutative Monoids, Kochel am See, October 1989, Report TUM-I9002, pages 1-10, TU M nchen, 1990.
- [CF69] P. Cartier and D. Foata, Probl mes Combinatoires de Commutation et de R arrangements, Volume 85 of Lecture Notes in Mathematics, Springer-Verlag, Berlin-Heidelberg-New York, 1969.
- [DR95] V. Diekert and G. Rozenberg, Editors, The Book of Traces, World Scientific, Singapore, 1995.
- [Fin01] O. Finkel, Topological Properties of Omega Context Free Languages, Theoretical Computer Science, Vol. 262 (1-2), July 2001, p. 669-697.
- [Fin03a] O. Finkel, Borel Hierarchy and Omega Context Free Languages, Theoretical Computer Science, Vol 290 (3), 2003, p. 1385-1405.
- [Fin03b] O. Finkel, Ambiguity in Omega Context Free Languages, Theoretical Computer Science, Volume 301 (1-3), 2003, p. 217-270.
- [Fin03c] O. Finkel, On the Topological Complexity of Infinitary Rational Relations, RAIRO-Theoretical Informatics and Applications, Volume 37 (2), 2003, p. 105-113.

- [Fin03d] O. Finkel, Undecidability of Topological and Arithmetical Properties of Infinitary Rational Relations, *RAIRO-Theoretical Informatics and Applications*, Volume 37 (2), 2003, p. 115-126.
- [FS03] O. Finkel and P. Simonnet, Topology and Ambiguity in Omega Context Free Languages, *Bulletin of the Belgian Mathematical Society*, Volume 10 (5), 2003, p. 707-722.
- [Fin04a] O. Finkel, On Recognizable Languages of Infinite Pictures, *International Journal of Foundations of Computer Science*, to appear.  
Available from <http://www.logique.jussieu.fr/www.finkel>
- [Fin04b] O. Finkel, An  $\omega$ -Power of a Finitary Language Which is a Borel Set of Infinite Rank, *Fundamenta Informaticae*, Volume 62 (3-4), 2004, p. 333-342.
- [Gas91] P. Gastin, Recognizable and Rational Languages of Finite and Infinite Traces, *Actes du STACS'91, Lecture Notes in Computer Science* 480, p. 89-104, 1991.
- [GP95] P. Gastin, A. Petit, Infinite Traces, Chapter in *The Book of Traces*, edited by V. Diekert et G. Rozenberg, World Scientific, p. 393-486, 1995.
- [GPZ94] P. Gastin, A. Petit, W. Zielonka, An Extension of Kleene's and Ochmanski's Theorems to Infinite Traces, *Fundamental Study, Theoretical Computer Science* 125, p. 167-204, 1994.
- [HM97] H.J. Hoogeboom and A. Muscholl, The Code Problem for Traces - Improving the Boundaries, *Theoretical Computer Science* 172, p. 309-321, 1997.
- [Kec95] A.S. Kechris, *Classical Descriptive Set Theory*, Springer-Verlag, 1995.
- [KK03] R. Kummets and D. Kuske, The Topology of Mazurkiewicz Traces, *Theoretical Computer Science*, Volume 305 (1-3), p. 237-258, 2003.
- [Kwi90] M. Kwiatkowska, A Metric for Traces, *Information Processing Letters*, 35, p. 129-135, 1990.
- [Lan69] L. H. Landweber, Decision Problems for  $\omega$ -Automata, *Math. Syst. Theory* 3 (1969) 4, 376-384.
- [Lec01] D. Lecomte, Sur les Ensembles de Phrases Infinies Constructibles a Partir d'un Dictionnaire sur un Alphabet Fini, *Séminaire d'Initiation a l'Analyse*, Volume 1, année 2001-2002.
- [LT94] H. Lescow and W. Thomas, Logical Specifications of Infinite Computations, In: "A Decade of Concurrency" (J. W. de Bakker et al., eds), Springer LNCS 803 (1994), 583-621.
- [Maz77] A. Mazurkiewicz, *Concurrent Programs Schemes and their Interpretation*, DAIMI Report PB-78, Aarhus University, Aarhus, 1977.
- [Mos80] Y. N. Moschovakis, *Descriptive Set Theory*, North-Holland, Amsterdam 1980.



- [Niw85] D. Niwinski, An example of Non Borel Set of Infinite Trees Recognizable by a Rabin Automaton, in Polish, Manuscript, University of Warsaw, 1985.
- [Niw90] D. Niwinski, Problem on  $\omega$ -Powers posed in the Proceedings of the 1990 Workshop “Logics and Recognizable Sets” (Univ. Kiel).
- [PP04] D. Perrin and J.-E. Pin, Infinite Words, Automata, Semigroups, Logic and Games, Volume 141 of Pure and Applied Mathematics, Elsevier, 2004.
- [Sim92] P. Simonnet, Automates et Théorie Descriptive, Ph. D. Thesis, Université Paris 7, March 1992.
- [Sim93] P. Simonnet, Automate d’ Arbres Infinis et Choix Borélien, C.R.A.S. Paris, t.316, Série 1, p. 97-100, 1993.
- [Sta86] L. Staiger, Hierarchies of Recursive  $\omega$ -Languages, Jour. Inform. Process. Cybernetics EIK 22 (1986) 5/6, 219-241.
- [Sta97a] L. Staiger,  $\omega$ -Languages, Chapter of the Handbook of Formal languages, Vol 3, edited by G. Rozenberg and A. Salomaa, Springer-Verlag, Berlin.
- [Sta97b] L. Staiger, On  $\omega$ -Power Languages, in New Trends in Formal Languages, Control, Cooperation, and Combinatorics, Lecture Notes in Computer Science 1218, Springer-Verlag, Berlin 1997, 377-393.
- [Tho90] W. Thomas, Automata on Infinite Objects, in: J. Van Leeuwen, ed., Handbook of Theoretical Computer Science, Vol. B ( Elsevier, Amsterdam, 1990 ), p. 133-191.